Models, representations, and simulations of plants and agro-ecosystems

Jean M. Thiéry
CEA/Cadarache, DSV-DEVM-Modélisation, 13108 Saint-Paul-lez-Durance, France

Summary: *Simulations* of reality are more and more developed for technical and scientific objectives and are based on *models* representing the different relevant phenomena. These models import data from databases and display their results with visible text or graphic representations. In simulations, reality is represented at three different levels: *computer representations* of previous measures stored into optimized databases, *visible representations* of model results and *mental representations* of scientists and observers. The importance of realistic visible representations is discussed with a landscape vegetation model for semi-arid ecosystems. Complete simulations of reality require the development of new representation techniques involving touch, smell, and taste and not only sight and hearing.

Models, *representations*, and *simulations* are often used interchangeably in many documents. In this article, we present different aspects of these three approaches, especially in the case of plants, agrosystems, and ecosystems.

DEFINITIONS OF MODELS, REPRESENTATIONS, AND SIMULATIONS

The three words *model*, *representation*, and *simulation* have various common meanings. Some of them apply directly to scientific disciplines: a *model* is a *"simplified description of a system to assist calculations and predictions"* (1) or a *"tentative description of a theory or system that accounts for all its known properties"* (2). A *representation* is something that represents, i.e., a *"description"*, a *"portrait"* or an *"equivalence"*, etc. (1, 2). It may imply an official equivalence in political or social life. A *simulation* is the *"act or process of having the appearance"* or an *"imitation"* or even a *"false appearance"* ! (2). Similar more or less ambivalent definitions can be found in Oxford dictionary (1) which adds another meaning: *"imitation of conditions of a situation with model, for convenience or training"*.

The Webster's definition of a *model* is more ambitious than the Oxford one, since a good model should be able to account for all known properties. This might be possible in some disciplines such as atomic physics. This is quite ambitious in agro-ecosystems, where the same object can be seen from different points of view and at different scales. For example, a clod of earth can be described by its contact to a farmer, by its mechanical properties to a civil engineer, by its organic matter content to a pedologist, by its microfauna to an ecologist, by its possible pollution to an environmentalist, etc.

Most classical models do not try to account for all known properties and are specialized by discipline for practical reasons: scientists cannot dominate many disciplines at the same time, team work is limited by human communication, and this tendency is unfortunately reinforced by the exponential growth of scientific knowledge. *Discipline models* trying to synthesize all information of a discipline may be too complex for every day research. They may be replaced by *simplified models* representing only the most important features of an actual research. These simplified models appear like *caricatures* or *cartoons*, while more *detailed models* would be like true

pictures. Detailed models can be as complicated as reality and cannot be tested easily. Simplified models lead to clearer predictions which can be easily tested by specific experiments. They play an important role in the analytic approach and can be used as elementary bricks for synthetic discipline models. Construction of synthetic models often uses the systemic approach which has been so successful in engineering: higher models are made of simpler models which can be easily tested separately.

The discipline approach is not satisfactory for fundamental reasons: mechanical, chemical and biological properties of an earth clod are strongly correlated and should be ideally described by a unique comprehensive model. The discipline approach is not adapted to most decisions which need a global point of view (e.g., daily irrigation decisions or long term soil conservation policies). Synthesis of different discipline models is not an easy task. Discipline models operate at different time or space scales (e.g. plant and microbe models). Some models may be quite elaborate while others are still crude by lack of information (this is already true in plant physiology where shoot models are much more sophisticated than root models). However, some multidisciplinary models with synthetic objectives have already been developed: e.g., the *EPIC* (Erosion Productivity Impact Calculator) model includes a crop growth component and computes water and wind erosion for different land management scenarios (3); the *ECOLECON* model simulates animal population dynamics and economic revenues in response to different forest landscape structure and timber management scenarios (4).

Everybody has his own *representation* of reality perceived by his senses (e.g., by *direct observation*). But different people may have different representations of the same reality (e.g., the same sound may be noise or music !). The scientific approach tries to obtain *objective representations* of reality, through the development of adequate concepts and measurements. Reality (or at least part of it) is represented by estimates (e.g., cloudy weather), scales (e.g., force 5 wind), or numbers (e.g., atmospheric pressure) whenever possible. These data may be stored in databases, which are *numerical representations* of reality. They can be printed in synthetic tables or displayed on curves, maps, images, or videos, which are *visible representations* of reality. When properly drawn, these visible representations should be objective enough so that different people should get the same *mental representation* of reality.

In scientific disciplines, *representation* and *simulation* may have the same meaning, i.e., a picture of reality. A *representation* may imply a better or more fundamental description of experimental facts but is often static like a portrait or a snapshot. A *simulation* has often a dynamic aspect, like in flight simulators.

The definition of a *simulation* as an *imitation* is akin to the definition of a picture. Ideally, a perfect simulation should have the same *visual appearance* as reality and an independent observer should not be able to distinguish between simulation and reality. Simulation should also present the same apparent *behaviour* as reality, flight simulators used for game or training (2nd Oxford definition). In practice, such simulations are the result of complex computations with detailed models for studied objects and their environment (lighting, shade, reflections, etc.). Quite realistic models for plant or landscape simulations have already been published and may be used for synthetic images (5).

Most simulations are less ambitious and only try to represent the evolution of selected variables of the system. The result is not an image of reality but a series of graphs displaying these variables versus time (chronicles) or versus one of them (trajectories) (6). Such *partial simulations* are also called *numerical experiments*, and sometimes *experiments* (!), which is quite misleading. In databases where interpolated data and computed results are mixed with pure experimental data, the origin of data should always be specified, for fundamental as well as statistical reasons. In particular, smoothed data should not be mixed with original experimental data since lost information cannot be retrieved (7, page 151).

Fig. 1 summarizes the main connections between *models, representations*, and *simulations*, in the case of a visual description of reality. For clarity, most of this article is presented as if sight was the unique sense involved in scientific research. Similar discussions apply to other senses (hearing, touch, smell, and taste), which may also play an important role in science, as explained in the following Conclusion on *a scientific multimedia representation of reality*.

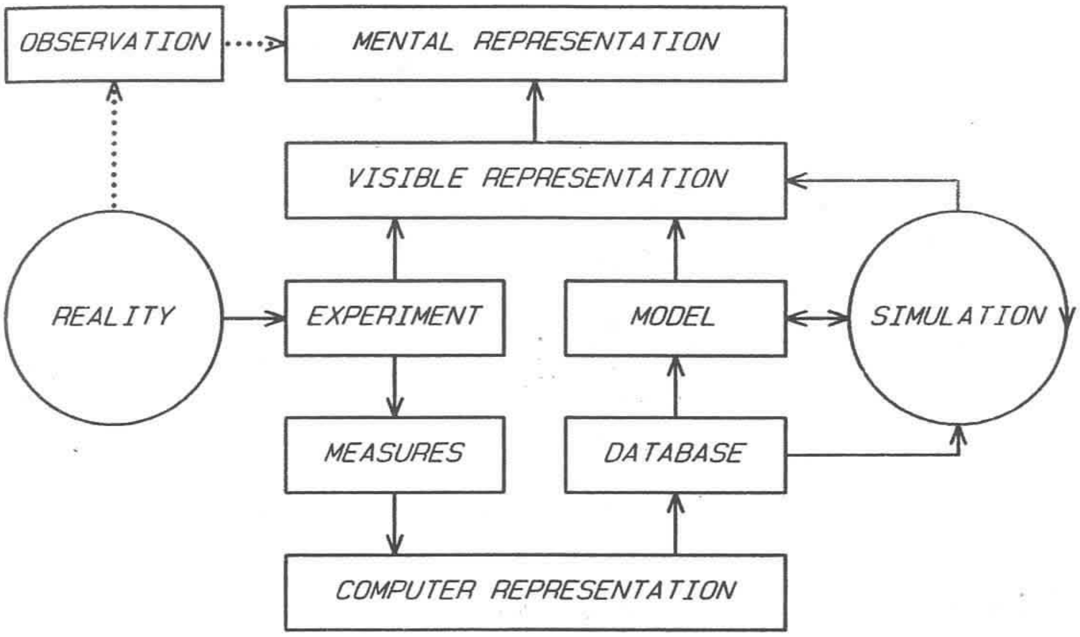

Figure 1: Main connections between *models*, *representations*, and *simulations*. This figure favours visible observations and representations: other senses, such as hearing, may also be involved (see the Conclusion on *a scientific multimedia representation of reality*).

Fig. 1 shows that *mental representations* of *reality* may be obtained by *direct observation*, which is possible in many sciences at human scale (agronomy, ecology, medicine, etc.). This direct approach may be useful for rapid diagnosis but may be quite subjective. Most scientific advances have been fostered by properly designed *experiments* producing objective *visible representations* (tables, graphs, images, or videos) which should induce the same *mental representations* on independent experts. *Measures* produced by experiments are more and more stored into *databases*, with *computer representations* allowing compact storage and fast retrieval. These databases are essential for most *models*, although some models need only estimated parameters and homogeneous or random initial conditions. Model results (like experimental results) can be highlighted by proper *visible representations* adapted to human understanding.

As explained in dictionaries, an observation may be either a *direct observation* (without any instrument) or a true *scientific observation*. On *Fig.1*, a *scientific observation* corresponds to the path going from *reality* to *mental representation* through *experiment* and *visual representation*.

Numerous psychological and technical researches try to clarify relations between *visible representations* and *mental representations*. These relations play an important role in computer ergonomics and in man-machine interactions, either in 2 dimensions (usual interfaces) or in 3 dimensions (virtual reality). They may be critical for the proper interpretation of model results. Independent tests may be necessary to check if different experts give identical interpretations of the same visible representation.

3-D visible representations are intensively used in *virtual reality* demonstrations. With fast workstations and appropriate equipment (glasses or helmet), everybody can visit historical monuments, painted caves, ..., or hostile environments. In agronomy, the same technology could be applied to botanical gardens, with possible seasonal changes within few minutes. Algorithms have already been tested for plant and flower representation (5) and have been used in synthetic images for architecture and films.

With modern computers and small databases, data can be stored in simple matrices or tables with no specific optimization. For larger databases, *numerical representations* should be optimized, with a classical trade-off between memory size and retrieval time. Most experimental results can be stored as real numbers on 4 bytes while others can be represented by simple integers on 2 bytes. Intermediate results have often to be stored on 8 bytes. Numerous codes have been developed for image files, where elementary points (pixels) can be stored on 1 up to 24 bits, with or without compression techniques (8). In *Geographical Information Systems* data can be stored either as matrices (raster mode) or as vectors defining river beds, communication lines or limits of homogeneous regions (9). Specific algorithms have been devised for fast retrieval of mixed data associating texts and numbers.

On a computer, a *model* is represented by a program and a *simple simulation* corresponds to a single run of this program working on a *database* and producing a *visible representation*. A *complete simulation* produces numerous *visible representations* obtained with the same *model* and different *data sets*. These different computations may be completely independent for statistical purposes or time correlated for *dynamic simulations*. The iterative aspect of *dynamic simulation* is represented on *Fig. 1* by an oriented circle importing data from the *database*, running the *model*, and displaying *visible representations* of results. In the case of a flight simulator,...

On *Fig 1*, *databases* and *models* are clearly separated, which is usually the case for numerical databases. In classical procedural programming, data and programs are completely independent so that the same program can be run with different data sets and the same data set can be used with different programs. This independence insures a better database integrity but may lead to some redundancy in programming. In *object-oriented databases*, part of the programming is included within the database itself and is available for all applications, e.g. object classes may include dedicated procedures (called methods) for usual data processing (such as specific corrections, filtering, interpolation, etc.).

In some disciplines, experiments cannot produce pure measures but only estimates or *qualitative results*, which can be written with conditional sentences *"IF ... THEN ..."*. These sentences may be included into expert systems, simulating the reasoning process of human experts. In that case, *Fig. 1* should be modified: *measures* should be replaced by *expert knowledge*, *computer representation* by an *artificial intelligence language* (usually *LISP* or *PROLOG*) and *database* and *model* should be merged into a *knowledge base*. More information on qualitative programming can be found in the associated article entitled *"Modeling the physical World"* (10), written with a different point of view. Faltings's article is centred on *model building processes* while this article is devoted to *information flows* between *reality* and *mental representations*.

EXAMPLES OF MODELS, REPRESENTATIONS, AND SIMULATIONS IN LANDSCAPE ECOLOGY

Examples of *models*, *representations*, and *simulations* can be found in any scientific disciplines, but in many cases simulations may be quite abstract since most phenomena can only be observed through complicated instrumentation. Many aspects of agro-ecosystems can still be directly observed, especially at the plant and landscape level, even if direct observation may be misleading (in landscapes, plastic bags may appear more polluting than disseminated biochemical products !).

The following discussion will be focused on a case study of *models*, *representations*, and *simulations* (with an article entitled "A *model simulating* the genesis of banded vegetation patterns in Niger" (11)). We have chosen the case of vegetation patterns which have been reported in many arid and semi-arid regions. In these patterns, vegetation tends to aggregate along contour lines into bands separated from each other by bare areas. We shall focus our attention on Sahelian vegetation patterns called *"brousses tigrées"* (12), i.e., *"tiger bushes"*. These bands can be easily observed or photographed from aeroplanes. They can also be perceived on the ground while driving through large

bare corridors (30 to 60 m wide) separating the vegetation bands (10 to 30 m wide).

Different hypotheses were published for the functioning of *"brousses tigrées"* and an informal model was presented by Ambouta (13). Such informal models, written in usual language without any mathematical formula, play a significant role in new sciences when researchers are looking for the main parameters describing the system. In his detailed description of *"brousses tigrées"*, Ambouta showed the important contribution of bare areas collecting water for vegetation bands with a large tree density.

The Ambouta model was recently formalized into the °*TIGREE*° landscape model based on cellular automata and depending only on two hypotheses which reflect *competition* and *synergy* between plants: the establishment, growth and survival of a given plant is affected *negatively* by plants situated up-slope and *positively* by lateral and down-slope plants (11). *Competition* is mostly a competition for water, while *synergy* encompasses many biological effects such as protection against predators or excessive transpiration, soil improvement by roots and termites, etc. Readers are referred to the original article for more details on the ecological aspects of *"brousses tigrées"*.

A landscape may sometimes be described by a simple aerial photograph, which could be considered as a good *representation*. For practical reasons, this representation may be simultaneously too detailed and not enough precise. Shepherds could be misled by photographs taken during the rainy season and showing temporary ponds. They could prefer traditional information sources. Car tracks could be confused with animal trails unsuitable for motor vehicles. Drivers could prefer simpler maps (or even

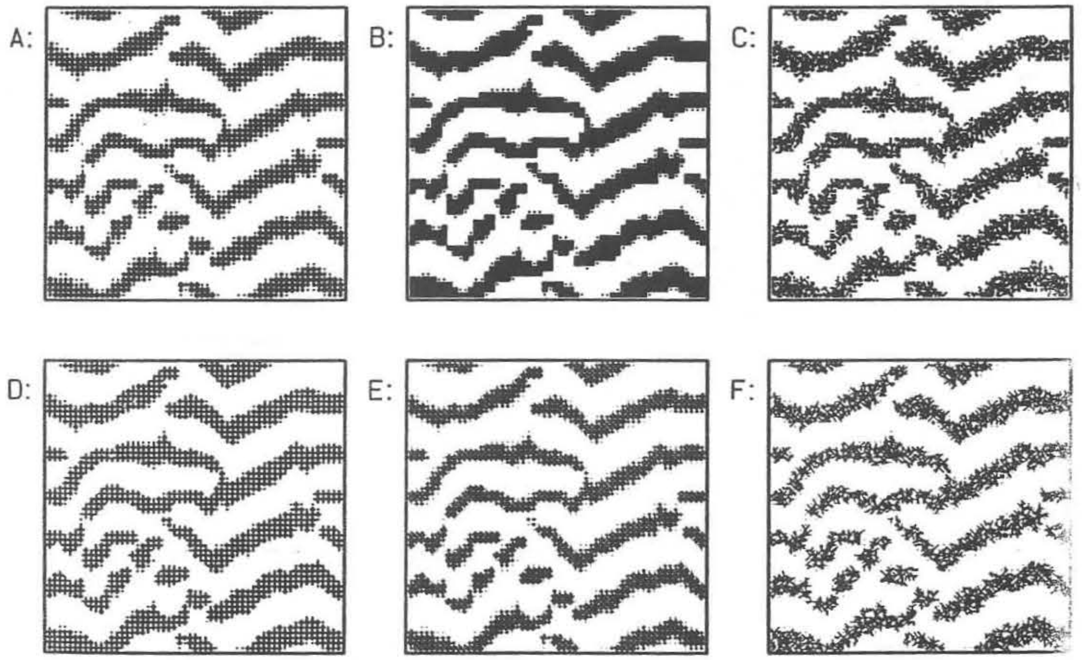

Figure 2: *Icons A* to *F* represent the same theoretical state matrix (*Fig. 3 Case E20*, in (11)). In each *icon*, sites are represented by specific symbols: dots (*A*), solid squares (*B*), randomly dotted squares (*C*), crosses (*D*), diamonds (*E*), random tree-like symbols (*F*). The size of geometrical symbols and the maximum number of elementary dots in random symbols are proportional to the state number.

sketches) which can be considered as specific *models* giving the essential information they need.

For the °*TIGREE*° model, the essential information of a *"brousse tigrée"* was vegetation and the landscape was formally divided into square sites (e.g., 5 m * 5 m) with vegetation classified into 4 states (*State 0* for bare areas, *State 1* for young or very stressed vegetation, *State 2* for intermediate vegetation and *State 3* for fully developed vegetation). The *"brousse tigrée"* is represented by a matrix S_{ij}, which is a compact *computer representation*, well adapted to computation and data storage (4 sites per data byte like in many graphics files) (8). This state matrix could be obtained by digitizing aerial pictures allowing the computation of proper vegetation indices. Experimental and theoretical state matrices could be compared by statistical methods adapted to landscape ecology. They can also be compared by direct visual comparison. *Fig. 2* displays different *visible representations*: *Icons A* to *F* correspond to the same theoretical state matrix and are fundamentally similar, e.g., with respect to topology. Their visual comparison points out the importance of an appropriate *representation*. Dots (*A*) and crosses (*D*) create unrealistic geometric gaps while solid squares (*B*), and to a lesser extent solid diamonds (*E*), induce unrealistic rectangular shapes. Randomly dotted squares (*C*) are more realistic but also induce some rectangular shapes. Random tree-like symbols (*F*) are the most realistic, probably because they are well adapted to the site description.

The introduction of random fluctuations for a better representation of reality is quite general in *synthetic images* and is the subject of numerous researches. Former synthetic images were too geometric and too regular to be credible. For artistic applications, where the public is the final judge, there is no objection against fluctuations in the pure tradition of impressionism. For scientific applications, fluctuations should be carefully controlled since they may introduce some hidden randomness into purely determinist models and influence the *mental representation*. Their effects could be risky if they were introduced during the intermediate steps of an iterative computation.

The °*TIGREE*° *model*, with its final *tree-like representation* was run in numerous *simulations* for parameter calibration. The published interaction matrix and its *competition* and *synergy* parameters were obtained by visual comparison with aerial photographs (11). Some management scenarios were also tested by *simulations* (14).

As previously explained, landscapes cannot be described by a unique photograph, map or model. Their complete *representation* uses more and more *Geographical Information Systems (GIS)* allowing the superposition of different points of view (15). For a typical *"brousse tigrée"*, an *"ideal GIS"* should contain at least the following maps: detailed vegetation distribution, village limits for wood managers and village councils, temporary ponds and trails for shepherds, micro-topography for hydrologists, car tracks for scientists and eco-tourists, etc. These maps should be scaleable in order to study specific vegetation bands as well as the whole banded system. They should be structured and give detailed information on specific objects (trees, termite mounds, etc.) ! Most real GIS are less ambitious and are focused on more specialized problems (15).

The previous case study of *models*, *representations*, and *simulations* was based on landscapes. Similar case studies could have been done on plant models, e.g., with the article entitled *"A model simulating above- and below-ground tree architecture with agroforestry applications"* (16). In this article, *AMAP*, a general *plant model* computes many realistic *visible representations* of trees.

VALIDATION OF SIMULATIONS

Validation of *partial simulations* (computing selected variables of the system) is now well established and is based on visual as well as statistical tests. Visual tests (with computed curves superimposed on corresponding data points) are generally quite efficient for the preliminary selection of appropriate models and scientists can often decide if the visual fit is realistic or not. Later on, proper statistical tests are necessary to choose the best fit of data by two appropriate models (or by the same model with different parameter values) (7, 17). A good fit is assumed when differences between computed and experimental values are completely random. Fit quality is expressed by classical tests such as root mean square deviations or X^2 (7). Most classical graphical or statistical tests can be run with standard software. In that case, the model

is written into a stand-alone program producing output files interpreted by a general purpose graphical and statistical package. This approach is well adapted to intensive computations run on large mainframe computers, while interpretations are done on user-friendly workstations or personal computers. Less demanding computations can be integrated into general purpose modelling tools allowing easier exchanges between computation and interpretation: e.g., *SIMULINK* (18), °*VOYONS*° (11, 14, 19).

Comprehensive simulations imitating reality are important in agro-ecosystems where many components (trees, landscapes, etc.) can be directly observed by men. Their validation may be quite difficult since *truth may be stranger than fiction* and the previous validation approach should be adapted. In general, validation is done by independent experts, choosing the best simulated representation or pointing out some unrealistic aspects (e.g., too large or too geometric leaves, etc.). Simulated objects (like trees) cannot be fitted directly to real ones and quantitative comparisons can only be indirect: proper statistical tests run separately on simulated and experimental objects should be coherent.

CONCLUSION: TOWARD A SCIENTIFIC MULTIMEDIA REPRESENTATION OF REALITY

Previous Sections were focused on the *visible representation* of reality. For centuries, scientific knowledge was mainly transmitted by printed media, with text and appropriate graphs. Oral transmission played an important role in teaching and in scientific discussions (in laboratories as well as in conferences) but this transmission could not be recorded until few decades. Numerous scientific films for the general public have already combined images, sounds and comments. They could be superseded by new multimedia documents allowing an interactive approach.

Multimedia documents, well adapted to the recording of natural sounds, have already fostered new researches. Bird songs may be used to pinpoint the limits of ecological niches and to look for changes due to pollution or agricultural practices (20). Animal cries are important for ethology, the scientific study of animal behaviour.

Other senses may be involved in technical or scientific perceptions of reality. Vibration perception by touch or hearing is crucial for sailing or engine tuning; sophisticated flight simulators should include simulations of accelerations and vibrations. Touch is also important in the precise control of computer devices such as mice and joysticks. Prostheses like artificial arms have to integrate proper pressure sensors. In agronomy, touch is often used by farmers and consumers for quality tests of edible products.

Smell and taste are important in chemistry and related sciences (biochemistry, pharmacology..., and gastronomy). Smell is already simulated by devices like smoke detectors, which unfortunately are not very specific. But, to our knowledge, models, like money, have no smell. Many new developments are necessary for an objective representation of reality for smell and taste.

Many robots may already be considered as *multimedia simulations* of animals or human beings, in the tradition of automata of the *XVIIIth* Century, although modern robots, working in hostile environment on Earth or on Mars and optimized for their task, have no resemblance with any living being. Conversely, for aesthetic and psychological reasons, prostheses should always be as good *simulations* as possible of human limbs.

ACKNOWLEDGEMENTS

The author thanks Dr Laurent COURNAC for a careful reading of the manuscript

REFERENCES

1. Oxford Dictionary 1979. The Concise Oxford Dictionary of Current English. Sykes, J. B. (Editor). Oxford University Press, Oxford, U. K.
2. Webster's Dictionary (1984). Webster's II New Riverside University Dictionary. The Riverside Publishing Company, USA.
3. Williams, J. R., Jones, C. A. and Dyke, P. T. 1984. Trans. ASAE. 27, 129.
4. Liu, J. 1993. Ecological Modelling. 70, 63.
5. Prusinkiewicz, P. and Lindenmayer, A. 1996. The algorithmic beauty of plants. Springer-Verlag, New York.
6. Brown, D. and Rothery, P. 1993. Models in Biology: Mathematics, Statistics and Computing. John Wiley & Sons, Chichester, U. K., Page 308.

7. Taupin, D. 1988. Probabilities, data reduction and error analysis in the physical sciences. Les Editions de Physique, Monographies de Physique, Les Ulis, France.
8. Murray, J. D. and van Ryper, W. 1994. Encyclopedia of Graphics File Formats. O'Reilly and Associates, Sebastopol, Cal.
9. Stow, D. A., 1993. In: Haines-Young, R., Green, D. R. and Cousins S. H. (Editors). Landscape Ecology and Geographical Information Systems. Taylor and Francis, London, U. K., Page 12.
10. Faltings, B. 1997. Modeling the physical World. In *Scientific representation n agronomy and plant physiology*. Couchat Ph. (Editor).
11. Thiéry, J. M., d'Herbès, J.-M. and Valentin, C. 1995. *J. Ecology*. 83, 497.
12. Clos-Arceduc, M. 1956. Bulletin de l'IFAN, série A. 7, 677.
13. Ambouta, K. 1984. Contribution à l'édaphologie de la brousse tigrée de l'ouest nigérien. Thèse de Docteur Ingénieur. Université de Nancy, France.
14. Thiéry, J. M., d'Herbès J.-M. and Valentin, C.. 1997. *"Modélisation de la réponse de brousses tigrées à différents modes de gestion"*. Accepted for publication in *Tendances nouvelles en modélisation pour l'Environnement*. Elsevier, Paris.
15. Haines-Young, R., Green, D. R. and Cousins, S. H. (Editors) 1993. Landscape Ecology and Geographical Information Systems. Taylor and Francis, London, U. K., Page 3.
16. de Reffye, P., Houllier, F., Blaise, F., Barthélémy, D., Dauzat, J. and Auclair, D. 1995. Agroforestry Systems, 30: 175.
17. Walter, E. and Pronzato, L. 1997. Identification of parametric models. Springer, London.
18. The Math Works, Inc. 1996. The Student Edition of SIMULINK. Prentice Hall, Englewood Cliffs, N. J.
19. Thiéry, J. M. 1991. *°VOYONS°*, programme de simulations conversationnelles en Physico-Chimie et en Agronomie. In *'Logiciels pour la Chimie'*, Antonot, N, Côme, G.-M., Gartiser, T., Guidon, J. and Soulié, E. (Editors). Soc. Fr. Chimie (Paris) et Agence Nat. Logiciel (CNRS, Nancy), Pages 292-293.
20. Balent, G. and Courtiade, B. 1992. Landscape Ecology. 6, 195.